THE STATE OF NEW HAMPSHIRE
Northern District of Hillsborough County

300 Chestnut Street
Manchester, NH 03101 2490
603 669-7410

APR 28 2009

MICHAEL G VALENTINE ESQ
OFFICE OF THE COUNTY ATTORNEY
300 CHESTNUT STREET
MANCHESTER NH 03101

- 07-E-0238 State of NH v. William Ploof

You are hereby notified that on April 28, 2009, the following order was entered in the above matter

re: DAUBERT HEARING:

(see copy of order attached hereto)

(Abramson, J.)

4/28/2009
Date

/s/ John Safford
Clerk of Court

cc: Marguerite Wageling, Esq.
Lisa Wolford, Esq.
Tony Sculimbrene, Esq.
Ross A. McLeod, Esq.

THE STATE OF NEW HAMPSHIRE

HILLSBOROUGH, SS. **SUPERIOR COURT**
NORTHERN DISTRICT

The State of New Hampshire

v.

William Ploof

No. 07-E-0238

ORDER

Respondent William Ploof ("respondent") is currently being held at the Secure Psychiatric Unit pending resolution of the State's petition to commit him for involuntary treatment under RSA chapter 135-E. Respondent pled guilty on February 23, 1998 to one count of Aggravated Felonious Sexual Assault and one count of Sexual Assault. He was sentenced to four to ten years at the New Hampshire State Prison, and twelve months at the House of Corrections. Although respondent's maximum release date under this sentence was June 15, 2007, he continues to be held pending resolution of the State's petition. See RSA 135-E:7, I (2005).

RSA chapter 135-E "create[s] a civil commitment procedure for the long-term care and treatment of sexually violent predators." RSA 135-E:1. RSA 135-E:3, I, provides that the commissioner of the New Hampshire Department of Health and Human Services ("DHHS") shall establish a multidisciplinary team ("MDT") "to evaluate whether persons convicted of ... sexually violent offense[s] who are eligible for release from total confinement meet the definition of a sexually violent predator." "[I]f there is an articulable basis to believe that the

person is likely to engage in acts of sexual violence," the county attorney or attorney general "may request that the multidisciplinary team assess and evaluate the person to determine whether the person is a sexually violent predator." RSA 135-E:3, III. "If the multidisciplinary team finds the person meets the definition of a sexually violent predator, the county attorney or attorney general may file a petition within 14 days with the superior court alleging that the person is a sexually violent predator and stating facts sufficient to support such allegation." RSA 135-E:6.

Pursuant to RSA 135-E:3, III, respondent was evaluated by an MDT to determine whether he was a candidate for RSA chapter 135-E commitment. The MDT was comprised of one DHHS employee and two psychologists experienced in the assessment and treatment of sex offenders. One of the psychologists was Dr. Eric Mart, who testified at the Daubert hearing in this matter. As part of its evaluation of respondent, the MDT was required to make a determination regarding respondent's risk of committing acts of sexual violence in the future. See RSA 135-E:6; RSA 135-E:2, XII (defining "sexually violent predator" as a person who "suffers from a mental abnormality or personality disorder that makes the person likely to engage in acts of sexual violence if not confined in a secure facility for long-term control, care, and treatment"). To make this determination, the MDT applied the Static-99, an actuarial tool that, based on the presence or absence of ten static factors, assesses an offender's relative recidivism risk as compared to other offenders. Essentially, an evaluator determines the presence or absence of the delineated factors in a particular individual's case, and scores

that individual according to the Static-99 scoring sheet. By conducting research on past populations of sexual offenders, the creators of the Static-99 have linked the Static-99 scores to particular risks of recidivism. (These recidivism risks, or norms, were recently updated: however, because this update occurred after the MDT's evaluation of respondent, the MDT did not consider these "new norms" in assessing respondent's likelihood of recidivism.) Once the MDT determined what recidivism risk correlated to respondent's score on the Static-99, it considered six additional factors it concluded were relevant to an assessment of respondent's recidivism risk. Based on all of this information, the MDT determined that respondent met the RSA 135-E:2, XII definition of a sexually violent predator.

On November 19, 2008, respondent requested a Daubert hearing, arguing that all evidence related to the MDT evaluation of respondent, including his score on the Static-99, should be excluded from trial because it does not comply with the requirements of RSA 516:29-a (2007). See Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 592-94 (1993). The Court held a three-day hearing on the matter. After consideration of the pleadings, the testimony elicited and arguments made at the Daubert hearing, and the applicable law, the Court finds and rules as follows.

The law governing the admissibility of evidence related to the MDT evaluation of respondent is set forth in RSA 516:29-a, which provides,

I. A witness shall not be allowed to offer expert testimony unless the court finds:

(a) Such testimony is based upon sufficient facts or data;

(b) Such testimony is the product of reliable principles and methods; and

(c) The witness has applied the principles and methods reliably to the facts of the case.

II. (a) In evaluating the basis for proffered expert testimony, the court shall consider, if appropriate to the circumstances, whether the expert's opinions were supported by theories or techniques that:

(1) Have been or can be tested;

(2) Have been subjected to peer review and publication;

(3) Have a known or potential rate of error; and

(4) Are generally accepted in the appropriate scientific literature.

(b) In making its findings, the court may consider other factors specific to the proffered testimony.

In Baker Valley Lumber Co. v. Ingersoll-Rand Co., 148 N.H. 609, 614 (2002), the New Hampshire Supreme Court adopted the standard of reliability articulated in Daubert, 509 U.S. at 592-94, to evaluate expert testimony. The Baker Valley Court noted that the Daubert standard, "establish[es] a more flexible standard of reliability that places special emphasis on four factors: (1) whether a theory or technique can be (and has been) tested; (2) whether the theory or technique has been subjected to peer review and publication; (3) the known or potential rate of error of a particular technique; and (4) the Frye general acceptance test." 148 N.H. at 614 (quoting Daubert, 509 U.S. at 593-94)

(quotations omitted); see also RSA 516:29-a (paralleling factors set forth in Baker Valley and Daubert). The Court noted that these factors "are not a definitive checklist or test," and other factors may also be appropriate. Baker Valley , 148 N.H. at 616-17.

Although the rules of evidence do not apply in this proceeding, see RSA 135-E:10, RSA 516:29-a does apply. Because RSA 516:29-a "unambiguously codified the four Daubert factors" adopted in Baker Valley, cases interpreting and applying Baker Valley remain instructive in this proceeding. See Baxter v. Temple, 157 N.H. 280, 284 (2008).[1]

"The trial court functions only as a gatekeeper, ensuring a methodology's reliability before permitting the fact-finder to determine the weight and credibility to be afforded an expert's testimony." Baker Valley, 148 N.H. at 616. Further, "the focus of the reliability analysis must be solely on principles and methodology, not on the conclusions they generate." Id. at 615 (quotations and brackets omitted). "[O]bjections to the basis of an expert's opinion go to the weight to be accorded the opinion evidence, and not to its admissibility. The appropriate method of testing the basis of an expert's opinion is by cross-examination of the expert." Id. at 615-16 (citation and quotations omitted).

In addressing the admissibility of expert testimony, the New Hampshire Supreme Court has

> recognized that evidence does not have to be infallible to be admissible. If it is of aid to a judge or jury, its deficiencies or weaknesses are a matter of defense, which affect the weight of the

[1] Although the hearing held in this matter is properly characterized as an RSA 516:29-a hearing, because the parties have referred to this proceeding as a Daubert hearing, the Court will use the same terminology.

> evidence but do not determine its admissibility. Thus, provided the proffered scientific evidence rests on sound scientific reasoning or methodology and properly can be applied to the facts in issue, it meets the requirements of Rule 702, even if the conclusion is novel or controversial.

State v. Dahood, 148 N.H. 723, 728 (2002) (citations omitted); see also State v. Langill, 157 N.H. 77, 85 (2008) (noting that, like RSA 516:29-a, Rule 702 requires that testimony be "based upon sufficient facts or data" and "the product of reliable principles and methods"). Furthermore, reliability under RSA 516:29-a is not synonymous with correctness, but rather,

> it signifies a much lower standard, to wit, trustworthiness. The overall purpose of Rule 702 and RSA 516:29-a is simply to ensure that a fact-finder is presented with reliable and relevant evidence, not flawless evidence. Indeed, it would be unreasonable to interpret [RSA 516:29-a,] I(c) as requiring that a single flaw or even multiple flaws in an expert's application of a particular methodology in all instances renders inadmissible the expert's entire testimony.

Langill, 157 N.H. at 87-88 (citations omitted).

In an RSA chapter 135-E trial, the factfinder is not charged with determining whether the particular respondent will recidivate, but whether that respondent is likely to recidivate. See RSA 135-E:2, XII (defining "sexually violent predator" as a person who "suffers from a mental abnormality or personality disorder that makes the person likely to engage in acts of sexual violence if not confined in a secure facility for long-term control, care, and treatment"); RSA 135-E:11. Thus, to comply with RSA 516:29-a, the expert testimony must reliably differentiate between classes of offenders that are likely to recidivate, and classes of offenders that are not likely to recidivate. The testimony does not have to reliably predict whether a particular individual will or

will not recidivate, and the science underlying the expert testimony need not have been developed for that purpose. Rather, the focus of an RSA 516:29-a analysis in this proceeding is whether the testimony and underlying science reliably differentiates between those "likely" to recidivate and those with a lesser recidivism risk. See RSA 135-E:2, XII.

Respondent alleges that several pieces of potential evidence do not comply with RSA 516:29-a. Specifically, respondent challenges: the science underlying the Static-99; testimony based on the Static-99's new norms; and the methodology used by the MDT in applying the Static-99, in considering several other factors not addressed by the Static-99, and in reaching an ultimate conclusion about respondent's recidivism risk. The Court will address each issue, in turn.

At hearing, the Court heard testimony from four expert witnesses about the validity and application of the Static-99, as well as the proper method for assessing an offender's recidivism risk. The Court first heard from Dr. Amy Phenix, the State's expert in forensic psychology. Dr. Phenix wrote the protocol for evaluating sexual offenders in California. In addition, she contributes to the website Static-99.org, and has drafted a proposed methodology for applying the Static-99 with the new norms. Dr. Phenix has testified for both the State and the respondent in sexual offender commitment proceedings. At hearing, respondent did not object to the State's Motion to qualify Dr. Phenix as an expert in forensic psychology, and the Court ruled that Dr. Phenix was so qualified.

The Court next heard testimony from Dr. Terence Campbell, the respondent's expert in forensic psychology. Dr. Campbell dedicates approximately one quarter of his practice in forensic psychology to the area of sexual offender recidivism. The majority of his work is conducted in the area of child sexual abuse. From 1993 to 1998, Dr. Campbell served as an editorial consultant for the publication "Issues in Child Abuse Accusations" published by the Institute of Psychological Therapies. Dr. Campbell has never testified for the State in sexual offender commitment proceedings. At hearing, the State did not object to respondent's Motion to qualify Dr. Campbell as an expert in forensic psychology, and the Court ruled that Dr. Campbell was so qualified.

The Court next heard testimony from Dr. Eric Mart, a member of the MDT that evaluated respondent. Dr. Mart testified about how he and the other members of the MDT evaluated respondent, and the basis for and appropriateness of their methodology. He also testified generally about the Static-99. At hearing, the State moved to have Dr. Mart qualified as an expert in clinical psychology. Although respondent at first attempted to limit Dr. Mart's expert qualification, respondent eventually agreed that Dr. Mart was qualified as an expert in clinical psychology, and the Court found that Dr. Mart was so qualified.

Lastly, the Court heard testimony from Dr. George Cobb, respondent's expert in applied statistics. Dr. Cobb testified about the statistical analysis involved with actuarial tools in general, and with the Static-99 in particular. Dr. Cobb also testified about the appropriateness of applying certain statistical

methods to data, as well as the appropriateness of applying statistical data to individuals. At hearing, the State did not object to respondent's Motion to qualify Dr. Cobb as an expert in applied statistics, and the Court found that Dr. Cobb was so qualified.

The Science Underlying the Original Static-99

The first question the Court must address is whether the science underlying the original Static-99 meets the standards of RSA 516:29-a. The Static-99, an actuarial tool developed in 1999 by R. Karl Hanson and David Thornton, is used to differentiate between individuals with higher or lower risks of sexual recidivism. The Static-99 has two main components: the test itself, including the factors listed and the method for scoring a particular sexual offender, and the corresponding recidivism risks or "norms." The test component consists of ten factors that could be present in a particular sexual offender. These factors are all empirically guided: that is, each factor has been statistically linked to an increase in sexual offender recidivism. The person administering the Static-99 determines the presence or absence of each potential factor regarding that particular offender, and scores him or her accordingly. Each score corresponds to a particular recidivism risk for individuals with that score. This corresponding recidivism risk is known in the field as a "norm."

A typical scoring worksheet for the test component of the Static-99 looks as follows:

	Risk Factor	Codes		Score
1	Young	Aged 25 or older Aged 18-24		0 1
2	Single	Ever lived with lover for at least two years? Yes No		 0 1
3	Index non-sexual violence	No Yes		0 1
4	Prior non-sexual violence	No Yes		0 1
5	Prior Sex Offenses	Charges None 1-2 3-5 6 +	Convictions None 1 2-3 4 +	 0 1 2 3
6	Prior Sentencing Dates (excluding index)	3 or less 4 or more		0 1
7	Any convictions for non-contact sex offenses	No Yes		0 1
8	Any unrelated victims	No Yes		0 1
9	Any stranger victims	No Yes		0 1
10	Any male victims	No Yes		0 1
	TOTAL SCORE			

When the Static-99 was first developed, Hanson and Thornton reported a set of original norms that corresponded to particular test scores. Dr. Phenix testified that the original norms were developed based on a population of 677 sexual offenders. Those norms have been cross-validated on many occasions. See State's Ex. 3. However, some claimed that the original norms reported an inflated recidivism risk, because they were based on older populations of sexual offenders. Recognizing that the overall rate of sexual offender recidivism was declining, in 2008, Hanson first notified the psychological community about the development of new norms. The new norms were generated based on two

modern populations of sexual offenders: a low risk population from the Correctional Service of Canada (the "CSC" population) and a high risk population. These populations included over 6,000 individuals. According to Dr. Phenix, the information used to generate the new norms came from a cross-validation of the Static-99. The new norms were first published in an article in the Association for Treatment of Sexual Abusers ("ATSA") forum. In the ATSA forum article, Hanson indicated that the old norms were no longer valid because they were based on outdated information from old populations of sexual offenders, and that they should not be used. The norms were the only aspect of the Static-99 that was revised.

The new norms differ from the old norms in two important ways. First, although the old norms set forth a specific recidivism risk relative to a particular score on the test component of the Static-99, the new norms establish a range of risks associated with each score. In addition, although the new norms are based on the same theory as the old norms, the new norms were generated using the statistical method of linear regression, as opposed to the survival analysis method used to generate the old norms. See State's Ex. 5., p. 4. By way of brief explanation, survival analysis involves simply reporting data about a particular population and using that data to generate probabilities, while logistic regression takes the data from the population and fits it to a particular logistic curve. See State's Ex. 5., p. 4. Because survival analysis includes only the data from that population, it cannot provide any further information. See State's Ex. 5., p. 4; see also Resp.'s Ex. B, p. 1. Furthermore, because survival analysis simply records

observed rates of recidivism and does not interpret the data, the old norms were generated based on, essentially, individual studies for each Static-99 score, and the norm associated with each score bore no relation to the norms set forth for the other scores. See State's Ex. 5., p. 4. This can result in random fluctuations in the data. See State's Ex. 5., p. 4. On the other hand, because logistic regression fits data to a particular logistic curve, the data for each score relates to the data for the other scores, thereby eliminating the "random fluctuations" that can occur in survival analysis. See State's Ex. 5., p. 4; see also Resp.'s Ex. B, p. 1.

At hearing, Dr. Cobb testified that the use of logistic regression, as opposed to survival analysis, did not invalidate the new norms. Indeed, he indicated that the use of logistic regression was an appropriate method as long as the data fit the curve. Dr. Cobb also testified that he was very experienced in applied statistics, and had read a great deal of literature related to that subject. Although he did not have sufficient information about the new norms to determine whether the data in fact fit the curve in this case, he stated that this did not invalidate the new norms.

Despite the changes to the norms, the reliability of the Static-99, as a whole instrument, has not been materially altered. Indeed, Dr. Cobb testified that the area under the curve ("AUC")[2] value of the new norms, a measure of the predictive accuracy of a particular tool, stayed "in the same ballpark" as the AUC value of the old norms. Dr. Cobb also noted that the AUC value of the Static-99

[2] Although Dr. Phenix generally referred to this measurement as an ROC value, she testified that the ROC value is the same thing as an AUC value, and that the difference was purely semantic in nature. For consistency, the Court will use "AUC" when referencing this concept.

indicated that it was "substantially better than guessing" at determining an offender's recidivism risk. Dr. Phenix testified that the AUC value for the old norms was approximately 0.68. The AUC value for the new norms is approximately 0.665. See Resp.'s Ex. B, p. 3.

Dr. Campbell and Dr. Cobb both expressed some concern about the basic theory behind the norms. They both testified that, because the norms are generated by using data from a particular population, they could only be reliably used to predict recidivism risks for individuals within that population. Because respondent was not a member of that population, they argued, neither the old nor the new norms could be reliably used to predict respondent's recidivism risk in this case. Despite this concern, neither Dr. Campbell nor Dr. Cobb testified that the information contained in the new norms was inaccurate: rather, they questioned the propriety of using that information to conclusively determine respondent's recidivism risk.

The experts raised several other issues regarding the Static-99. Although Dr. Phenix testified that actuarial tools like the Static-99 are used regularly to assess an offender's recidivism risk, and that such tools are generally accepted for that purpose, Dr. Phenix testified that it is recommended that evaluators use one or more actuarial tools in assessing a particular offender. Dr. Campbell testified that the Static-99 should not be used as a stand-alone instrument: that is, an evaluator should not base his or her opinion about an offender's recidivism risk solely on the Static-99. Dr. Phenix testified that, despite the general practice of using one or more actuarial tools in evaluating an offender, the use of multiple

actuarial tools did not increase the predictive accuracy of the assessment.

Dr. Cobb and Dr. Campbell both discussed the inter-rater reliability of the Static-99. By way of brief explanation, the inter-rater reliability measures whether a tool is subject to inconsistent applications: that is, how likely it is that two evaluators would apply the tool to the same individual and reach different conclusions. Although Dr. Cobb stated that the Static-99 had a relatively low inter-rater reliability, he based his conclusion on only one of fourteen studies addressing the inter-rater reliability of the Static-99. That one study reported an inter-rater reliability that was significantly lower than that reported in the other studies.

Based on the testimony and evidence offered at hearing, the Court finds that the test component of the Static-99 meets the requirements of RSA 516:29-a. First, the science underlying the test component of the Static-99 is based upon sufficient facts or data. See RSA 516:29-a, I(a). Both Dr. Phenix and Dr. Campbell testified that the original Static-99 was validated or developed using a population of 677 sexual offenders, and that it had been cross-validated on a number of subsequent unrelated populations. Furthermore, Dr. Phenix testified that each factor considered by the Static-99 was empirically validated. Based on this testimony, the Court finds that the test component of the Static-99 is based on sufficient facts or data to comply with RSA 516:29-a, I(a).

In addition, the Court finds that the test component of the Static-99 is the product of reliable principles and methods. See RSA 516:29-a, I(b). Both Dr. Phenix and Dr. Campbell testified that actuarial tools are commonly used in

forensic psychology, and that an actuarial tool such as the Static-99 has an accuracy rate that is significantly better than chance. Although Dr. Campbell indicated that the Static-99 should not be used alone in assessing an offender's recidivism risk, this does not mean that the test component of the Static-99 is unreliable. Rather, it impacts the reliability of an assessment based solely on Static-99. Furthermore, the method for scoring the Static-99 is reliable. Rather than using clinical judgment, a practice both Dr. Phenix and Dr. Campbell described as having an accuracy rate that was only slightly better than chance, the Static-99 is scored based on the presence or absence of objective factors. Dr. Cobb testified that the Static-99 has a low inter-rater reliability, meaning that there is a risk that two professionals scoring the same individual would give that individual different Static-99 scores. However, as previously noted, Dr. Cobb based his conclusion on only one of fourteen studies addressing the inter-rater reliability of the Static-99, and that one study reported an inter-rater reliability that was significantly lower than that reported in the other studies. Accordingly, the Court does not credit Dr. Cobb's testimony on this issue. Furthermore, Dr. Cobb agreed that the Static-99 was substantially better than chance or guesswork at predicting sexual offender recidivism. Based on the evidence presented, the Court finds that the test component of the Static-99 is the product of reliable principles and methods, and that it therefore complies with RSA 516:29-a, I(b).

RSA 516:29-a, I(c) relates to how the principles and methods underlying the Static-99 were applied in a particular case. Accordingly, the Court will address this component of the statute when addressing the MDT evaluation of

respondent (below).

The Court will now consider the test component of the Static-99 under RSA 516:29-a, II. Both Dr. Phenix and Dr. Campbell agreed that the test component of the Static-99 has been tested. See RSA 516:29-a, II(a)(1). Indeed, both experts discussed the extensive cross-validation conducted on this tool. Furthermore, an individual's score on the Static-99 can be tested because the factors listed in the Static-99 are objective and, for the most part, static. That is, an individual's score on the Static-99 is unlikely to change over time. Dr. Phenix testified that an individual's score on the Static-99 would change over time only if an offender went from being below 25 years of age to above 25 years of age, lived with a significant other for at least two years, or reoffended. Because an individual's score on the Static-99 is unlikely to change over time, it is possible to test the validity of a particular application of the Static-99 at some later date. Based on the evidence presented, the Court finds that the test component of the Static-99 complies with RSA 516:29-a, II(a)(1).

In addition, both Dr. Phenix and Dr. Campbell testified that the test component of the Static-99 has been subjected to peer review and publication. See RSA 516:29-a, II(a)(2). Indeed, at hearing, the parties referenced a number of articles and studies about the Static-99. Accordingly, the Court finds that the test component of the Static-99 complies with RSA 516:29-a, II(a)(2).

Furthermore, the test component of the Static-99 has a known or potential rate of error. See RSA 516:29-a, II(a)(3). During the process of cross-validation, the Static-99 has been found to have an AUC value between 0.71 and 0.68. As

testified to by Dr. Phenix and eventually on cross-examination by Dr. Campbell, this means that the Static-99 can be described as having "moderate" predictive accuracy. In addition, the inter-rater reliability of the Static-99 has been studied on a number of occasions, and most of those studies indicate that the Static-99 has moderate to high inter-rater reliability. Based on the foregoing, the Court finds that the test component of the Static-99 complies with RSA 516:29-a, II(a)(3).

Finally, the test component of the Static-99 is generally accepted in the appropriate scientific literature. See RSA 516:29-a, II(a)(4). Dr. Phenix testified that the Static-99 is generally accepted in the field of forensic psychology, and that a number of states mandate use of the Static-99 when evaluating sexual offenders. In addition, Dr. Phenix testified that the use of actuarial tools, generally, has been accepted in the relevant scientific literature. Furthermore, a number of courts have considered this issue and found that actuarial risk assessments like the Static-99 are generally accepted in the relevant scientific literature, and these findings, though not dispositive, create an inference that the Static-99 has such acceptance. See, e.g., In re Commitment of Simons, 213 Ill.2d 523, 535-36 (2004) ("[T]here is no question that [actuarial risk assessment] is generally accepted by professionals who assess sexually violent offenders and therefore is perfectly admissible in a court of law. As of this writing, experts in at least 19 other states rely upon actuarial risk assessment in forming their opinions on sex offenders' risks of recidivism. See State ex rel. Romley v. Fields, 35 P.3d 82, 89 (Ariz. 2001); People v. Therrian, 6 Cal.Rptr.3d 415, 419-20 (2003);

Roeling v. State, 880 So.2d 1234, 1238-40 (Fla. App. 2004); In re Detention of Holtz, 653 N.W.2d 613, 619 (Iowa App. 2002); In re Care & Treatment of Teer, No. 89,652, 90 P.3d 379, slip op. at 3-4 (Kan. App. 2004) (unpublished order); Commonwealth v. Wright, No. 032449A, 2004 WL 1690388, slip op. at 1 (Mass. Super. 2004); In re Risk Level Determination of R.B.P., 640 N.W.2d 351, 353-56 (Minn. App. 2002); Goddard v. State, No. 25779, 144 S.W.3d 848, 850-51 (Mo. App. 2004); State v. Legg, 84 P.3d 648, 651 (Mont. 2004); Slansky v. Nebraska State Patrol, 685 N.W.2d 335, 345-49 (Neb. 2004); In re Commitment of R.S., 801 A.2d 219, 220-21 (N.J. 2002); People v. Girup, 9 A.D.3d 913 (N.Y. 2004) (mem. op.); In re D.V.A., 676 N.W.2d 776, 778-80 (N.D. 2004); State v. McKinniss, 795 N.E.2d 160, 165-66 (Ohio 2003); State v. Gibson, 66 P.3d 560, 564-65 (Or. 2003); In re Care & Treatment of Tucker, 578 S.E.2d 719, 721 (S.C. 2003); In re Commitment of Morales, 98 S.W.3d 288, 291 (Tex. Ct. App. 2003); In re Detention of Thorell, 72 P.3d 708, 724-25 (Wash. 2003); In re Commitment of Tainter, 655 N.W.2d 538, 544 (Wis. 2002) . . . Eight of these states have directly addressed the Frye [general acceptance] question and concluded either that Frye is inapplicable to actuarial risk assessment [because the general acceptance question applies only to new or novel scientific methodologies] or that actuarial risk assessment satisfies the general acceptance standard."). Although Dr. Campbell testified that the Static-99 is not generally accepted, the Court does not credit that testimony as it is outweighed by contrary evidence. Further, the fact that Hanson has discontinued the original norms does not alter the generally accepted status of the test component of the Static-99, because the

norms are not involved in the test component, but are instead a subsequent step to the test component. Based on the evidence presented, the Court finds that the test component of the Static-99 complies with RSA 516:29-a, II(a)(4). Accordingly, the members of the MDT may testify about their testing of respondent using the Static-99, and about the numeric score (based on the Static-99 scale of 0 to 12) respondent received on the Static-99.

In light of the foregoing, the Court finds that the test component of the Static-99 complies with the requirements of RSA 516:29-a. The remaining issue with regard to the original Static-99 is whether the original norms comply with RSA 516:29-a. The Court finds that they do not. At hearing, both Dr. Phenix and Dr. Campbell testified that, consistent with Hanson's directive, the original norms of the Static-99 should be replaced with the new norms. This testimony essentially established that the original norms do not comply with RSA 516:29-a because they are no longer considered reliable by experts in the field and are not accepted in the applicable scientific literature. Accordingly, the Court finds that the original norms of the Static-99 do not comply with RSA 516:29-a, and are therefore inadmissible at the trial in this matter. Thus, testimony about how respondent's score on the Static-99 would correspond to a recidivism risk under the old norms is inadmissible. Likewise, testimony that is based on the old norms is inadmissible.

The Ranges of Recidivism Risks Set Forth in the New Norms

Based on the evidence presented at the Daubert hearing, the Court finds that the science underlying the new norms of the Static-99 meets the

requirements of RSA 516:29-a. As previously noted, the new norms were first introduced at a conference in 2008, and were ultimately published in the ATSA forum article in January of 2009. The new norms have also been published on the Static-99.org website. The new norms were developed based on studies of two recent populations of sexual offenders. Hanson developed the new norms using these newer populations in recognition of the fact that recidivism rates have been declining over time. State's Ex. 5 at p. 1. "Compared to the original norms, the new norms are based on more offenders, more complete data, and more recent, representative samples." State's Ex. 5 at p. 2.

Although derived using similar methodology, the new norms have one major difference from the original norms: rather than reporting a single percentile for recidivism risk, the new norms provide a range of risk. These ranges were derived based on two distinct portions of the relevant population: the CSC samples, which represent lower risk offenders, and the pre-selected high risk samples. Hanson recognized that the change from a single recidivism risk to a range of risks posed new challenges for evaluators, and made the following recommendation to assist them in utilizing the new norms:

> Currently, our recommendation is to report recidivism estimates with the new norms in two stages. The first stage involves reporting an empirically-derived range of recidivism risk. The recidivism estimates from the CSC samples represent the lower bound of the range and the preselected high-risk samples are the upper bound of the range . . . The second stage involves making a professional judgment as to where a particular offender is likely to fall within that range. This judgment represents a separate task from reporting the empirical recidivism rates; currently, there is no research to assess how well evaluators are able to make this judgment.

State's Ex. 5 at p. 3. Hanson explained that, like the original Static-99, the new norms provide an indication of an offender's recidivism risk relative to other sexual offenders. State's Ex. 5 at p. 5. Hanson also noted that, although the new norms were ready for use, they would be updated over time in order to respond to any new patterns in sexual offender recidivism. State's Ex. 5 at p. 6.

Based on the testimony and evidence offered at the hearing, the Court finds that the Static-99 new norms[3] meet the requirements of RSA 516:29-a. First, the Court finds that the new norms are based upon sufficient facts or data. See RSA 516:29-a, I(a). Although the population used to generate the original norms included only 667 sexual offenders, the new norms were generated based on a population of over 6,000 sexual offenders. In addition, when compared to the original norms, the new norms appear to recognize the overall decline in sexual offender recidivism: for Static-99 scores that were specifically addressed by the original norms[4], the new norms, overall, show a lower recidivism rate than that reported by the original norms for the 5 and 10 year periods. Indeed, for offenders with Static-99 scores of 4 or more, the "upper bounds" of the ranges reported by the new norms are lower than the corresponding recidivism rates under the original norms. Compare State's Ex. 5, p. 3, with State's Ex. 2. Based on the large population used to generate the new norms, and the fact that the new norms appear consistent with the overall trend of declining sexual offender

[3] In using the phrase "new norms," the Court refers to the recidivism ranges reported by the new norms. "New norms" does not refer to the assignment of a particular recidivism risk within those ranges to respondent.

[4] Another change from the original norms to the new norms is that the original norms provided a single recidivism risk for offenders with a Static-99 score of 6 or higher, while the new norms provide recidivism ranges specific to scores of 0 through 9, and then provide a recidivism risk for scores of 10 or higher.

recidivism, the Court finds that the new norms meet the requirements of RSA 516:29-a, I(a).

In addition, the Court finds that the new norms are the product of reliable principles and methods. See RSA 516:29-a, I(b). In developing the new norms, Hanson

> used logistic regression analysis to calculate recidivism estimates. In simplest terms, regression produces a "line of best fit" that models the relationship between an independent variable (Static-99 scores) and a dependant variable (the probability of recidivism). The slope of the line tells us the average increase in the probability of recidivism associated with each one-score increase on Static-99. The intercept of the line (where the line intersects with the y-axis) tells us the predicted recidivism rate for a Static-99 score of 0. Using both the intercept and the slope, regression allows us to predict recidivism rates for any score on Static-99. Logistic regression is a specific form of regression that transforms the dependent variable (i.e., the odds of recidivism) into its natural logarithm and is more appropriate for use with dichotomous outcome variables (i.e., recidivism).
>
> An advantage of logistic regression is that it uses information on the relationship between Static-99 and recidivism in the full dataset to make predictions for a given score. This eliminates the logical anomaly whereby offenders with a certain score can have slightly higher estimated recidivism rates than offenders with a higher score. In other words, it smoothes out the random fluctuations inherent in survival analysis and likely provides better estimates of the "true" relationship between the variables. Logistic regression is appropriate to use for generating recidivism estimates as long as the data approximate a logistic distribution (this assumption is satisfied [by the new norms]).

State's Ex. 5, p. 4; see also State's Ex. 5, p. 5, Figure 2 (illustrating that logistic regression, the method used to calculate the new norms, and survival analysis, the method used to calculate the original norms, produce generally consistent results, but that logistic regression avoids the random fluctuations seen in survival analysis). At hearing, the Court heard no testimony regarding the

advantages or disadvantages of using logistic regression versus survival analysis. Further, Dr. Cobb could not opine whether logistic regression was appropriate in this case because he had not seen the relevant data. He indicated that, based on what he knew, the use of logistic regression did not invalidate the new norms. Thus, based on the evidence before the Court, the Court finds that the new norms comply with RSA 516:29-a, I(b).

RSA 516:29-a, I(c) relates to how the new norms were applied in a particular case. Accordingly, the Court will address this component of the statute when addressing the MDT evaluation of respondent (below).

The Court will now address the admissibility of the new norms under RSA 516:29-a, II. Consistent with the mandates of RSA 516:29-a, in conducting this analysis, the Court considers whether the "theories or techniques" used to generate the new norms comply with RSA 516:29-a, II, and not whether the new norms themselves meet the requirements of RSA 516:29-a, II. See Baker Valley, 148 N.H. at 615 (noting that the court's focus in addressing the admissibility of expert testimony is on the methodology used, and not on any conclusions generated thereby). First, although the new norms have not yet been cross-validated, Dr. Campbell testified that they can be tested. See RSA 516:29-a, II(a)(1). Indeed, according to the testimony of both Dr. Phenix and Dr. Campbell, the old norms were tested on a number of occasions. This indicates that it is possible to test the accuracy of a recidivism risk derived from group data. In addition, Dr. Campbell testified that it is possible to test the new norms. That is, it is possible to test the accuracy of the information obtained through the use of

logistic regression. Based on this testimony, the Court finds that the theories and techniques underlying the new norms comply with RSA 516:29-a, II(a)(1).

In addition, the theories and techniques underlying the new norms have been subjected to peer review and publication. See RSA 516:29-a, II(a)(2). The old norms were subjected to extensive peer review and publication, as established by the numerous articles both parties referenced at the hearing in this matter. Thus, the theory that group data can be used to generate norms has been subjected to peer review and publication. In addition, the new norms were published in the ATSA forum article. Further, the new norms have been reviewed by other members of the profession. For example, Dr. Phenix has written a protocol for use in California regarding how to use the new norms. In addition, Dr. Campbell has had an opportunity to assess the new norms and provide any criticisms he may have. Additionally, the technique of logistic regression has been subjected to peer review. Indeed, Dr. Cobb was able to explain what statisticians look for in assessing whether logistic regression is appropriate in a particular case. Based on the evidence before the Court, the theories and techniques underlying the new norms comply with RSA 516:29-a, II(a)(2).

Further, the Court finds that the new norms have a known or potential rate of error. See RSA 516:29-a, II(a)(3). Dr. Phenix indicated that the new norms could be cross-validated. Likewise, in noting that the new norms had not yet been cross-validated, Dr. Campbell conceded that such cross-validation was possible. In addition, the new norms have an AUC value of 0.665, see Resp.'s

Ex. B, p. 3, and, as such, have a known rate of error. Thus, the new norms comply with RSA 516:29-a, II(a)(3).

The Court also finds that the theories and techniques underlying the new norms are generally accepted in the appropriate scientific literature. See RSA 516:29-a, II(a)(4). First, Dr. Phenix testified that the old norms were generally accepted. Indeed, many courts that addressed the original Static-99 determined that it enjoyed such general acceptance. See, e.g., In re Commitment of Simons, 213 Ill.2d at 535-36. Thus, the theory that group data can be used to generate norms has been generally accepted. The remaining issue, then, is whether the technique of logistic regression is generally accepted under RSA 516:29-a, II(a)(4). Dr. Phenix testified that the new norms are generally accepted. In addition, her protocol for California's evaluation of sexual offenders utilizes the new norms. Furthermore, Dr. Cobb testified that he had reviewed a great deal of literature relating to the Static-99, and to actuarial tools in general. Dr. Cobb also testified that logistic regression may be properly applied to data such as the group data used to generate the new norms as long as that data "fits the curve" used in logistic regression. This testimony further indicates that the technique of logistic regression is a generally accepted technique. Although Dr. Campbell testified that the new norms were not generally accepted, the Court finds that this testimony is outweighed by contrary evidence. Based on the evidence before the Court, the Court finds that the theories and techniques underlying the new norms comply with RSA 516:29-a, II(a)(4).

There are two additional factors the Court finds relevant to the

admissibility of the new norms. See RSA 516:29-a, II(b). First, the evidence establishes that the new norms are an improvement over the original norms, which have been admitted into evidence in numerous jurisdictions. See, e.g., In re Commitment of Simons, 213 Ill.2d at 535-36. Dr. Phenix testified that the new norms were an improvement. In addition, the new norms are better than the original norms because they provide a range of risk rather than a single number, which permits evaluators to consider each offender as an individual and determine whether that individual is more like a CSC or low risk offender or a high risk offender. Further, the new norms are based on more contemporary data. Indeed, Dr. Campbell's testimony regarding the new norms indicated only that they had not yet been cross-validated, but expressed no negative opinion about the ranges reported in the new norms. These considerations support introducing this evidence at the trial in this matter.

In addition, the Court notes that evidence like this is particularly helpful to a jury in this sort of case because it helps the jury differentiate between sexual offenders and determine whether a particular respondent is "likely to engage in acts of sexual violence if not confined in a secure facility." See RSA 135-E:2, XII. Because the new norms provide the jury with a framework for making this determination, but do not assign a particular recidivism risk to an individual, the Court finds that such evidence will be particularly helpful to the jury in this case. In other words, this evidence gives the jury some guidance, without making an ultimate conclusion about respondent's recidivism risk. This consideration further supports admission of this evidence at trial.

In light of the foregoing, the Court finds that the new norms of the Static-99 comply with the requirements of RSA 516:29-a. The Court will now address the admissibility of testimony regarding where respondent's recidivism risk falls within the range established by the new norms. As noted in the ATSA forum article, use of the new norms to report an individual offender's recidivism risk is best done in two stages: first, report the empirically-derived range of recidivism risk set forth in the new norms, and then make "a professional judgment as to where a particular offender is likely to fall within that range." State's Ex. 5, p. 3. The ATSA forum article acknowledges that there is "no research to assess how well evaluators are able to make this judgment." State's Ex. 5, p. 3. Dr. Phenix and Dr. Campbell both testified about the inaccuracies associated with clinical judgment, a necessary step for an expert to place respondent at a particular risk within the new norms. Based on the evidence before the Court, such use of clinical judgment would not meet the RSA 516:29-a, I(b) requirement that testimony be based on a reliable method. Accordingly, no expert may testify about where respondent's recidivism risk falls within the range set forth by the new norms, nor may an expert provide testimony that is based on having selected respondent's recidivism risk from the range set forth by the new norms.

The Evaluation By and Conclusions Of the MDT

Based on the evidence presented at the Daubert hearing, the Court finds that the evaluation by and conclusions of the MDT partially meet the requirements of RSA 516:29-a. There are three main components of this evidence: the MDT's application of the Static-99 with respect to respondent, the

other factors the MDT considered in determining respondent's recidivism risk, and the MDT's final conclusion about respondent's recidivism risk. The Court will address each component, in turn.

Because the Court has previously addressed the test component of the Static-99 as it relates to RSA 516:29-a, I(a)-(b) and RSA 516:29-II, the Court need only address whether the MDT reliably applied the test component of the Static-99 in this case. See RSA 516:29-a, I(c). At hearing, Dr. Mart testified that he and another member of the MDT each scored respondent on the Static-99. Although this scoring was not done on the scoring form, both reviewed the form and calculated respondent's score individually. Dr. Mart then compared his proposed score for respondent with that of the other MDT member, and noted that they had assigned respondent the same score. This application of the test component of the Static-99 permitted the MDT to check each MDT member's scoring of respondent. Because both MDT members arrived at the same score individually, the Court finds that this is evidence of a reliable application of the test component of the Static-99. See RSA 516:29-a, I(c).

At hearing, Dr. Cobb testified that the Static-99 could not be applied to an individual because the new norms are based on group data, and respondent was not a part of that group. Thus, Dr. Cobb testified, the new norms do not provide respondent's recidivism risk. However, this distinction does not impact the admissibility of the test results. Rather, it indicates only that experts must be careful in describing the new norms. The new norms do not report a particular individual's recidivism risk. The proper characterization of the new norms is that

individuals with a particular score on the Static-99 have been found to have a recidivism risk within the range set forth in the new norms. In other words, experts may testify about respondent's score on the one to twelve scale of the Static-99, and that offenders with that score have been found to recidivate at the rate set forth in the ranges of the new norms. See U.S. v. Shields, No. 07-12056-PBS, 2008 WL 544940, at *1 (D. Mass. Feb. 26, 2008) ("[T]he government shall not use the term 'high risk' in reference to a score on a Static-99 test without explaining its meaning in statistical terms (e.g., 'Respondent scored a[n 8] on the Static-99. Individuals with this score have been found to re-offend at [the range set forth in the new norms]. These percentages are group estimates; they do not directly correspond to the recidivism risk of an individual offender.')"). Subject to this clarification, the MDT may testify about respondent's score on the Static-99, and the recidivism range set forth in the new norms that corresponds to that score.

The Court will now address the "other factors" the MDT used to assign respondent a particular recidivism risk. These other factors included: (1) that respondent has never successfully completed treatment for his sexual behavior, (2) that respondent has not shown remorse, (3) that respondent has not shown empathy for his victims, (4) that respondent has admitted to victimizing between twenty and fifty individuals, (5) that respondent is a chronic offender of a predatory nature, and (6) that respondent's victims were male. The Court will address the admissibility of each factor.

The first factor the MDT considered in determining respondent's recidivism risk was that respondent has never successfully completed treatment for his sexual behavior. At hearing, Dr. Mart testified that this factor is empirically validated. In other words, studies have shown a link between failure to complete treatment and higher recidivism rates for sexual offenders. Furthermore, Dr. Mart noted that the Static-99 does not consider an individual's failure to complete treatment. Dr. Phenix and Dr. Campbell agreed that the use of empirically-guided clinical judgment to adjust the results of an actuarial tool may decrease the overall predictive accuracy of a particular assessment. However, this concern, at most, indicates that the MDT members should not be permitted to testify about how this factor should be used to assign respondent a particular recidivism risk within the range set forth by the new norms. It does not indicate that the jury should not hear evidence regarding this empirically-guided factor. Indeed, Dr. Phenix testified that the factors set forth in the Static-99 are all empirically-guided factors, and that the use of empirically-guided factors to assess an individual's recidivism risk is appropriate. Based on the evidence before the Court, the Court finds that this factor, standing alone, complies with the requirements of RSA 516:29-a. Thus, any expert, including a member of the MDT, may testify that an individual's failure to complete treatment has been linked to increased recidivism rates. Further, the State may introduce evidence that respondent has failed to complete treatment. However, because it would be based on unreliable clinical judgment (according to Dr. Phenix and Dr. Campbell), no expert may quantify the impact that this factor's presence has on

respondent's recidivism risk at trial. In other words, an expert may testify that the presence of this factor should be considered, but may not tell the jury how much weight to give it.

In determining respondent's recidivism risk, the MDT also considered that respondent has not shown remorse and has not shown empathy for his victims. At hearing, Dr. Mart admitted that these factors were not empirically guided, but that their consideration just made "good logical sense." Because these factors are not empirically guided, testimony about these factors in relation to respondent would not be based on sufficient facts or data. See RSA 516:29-a, I(a). Furthermore, reliance on such "logic" introduces the MDT's clinical judgment into its evaluation of respondent. As Dr. Phenix and Dr. Campbell testified, clinical judgment has a level of predictive accuracy that is only slightly better than chance. Accordingly, testimony about these factors would not be the product of reliable principles or methods. See RSA 516:29-a, I(b). Because it would not comply with RSA 516:29-a, I(a)-(b), testimony about respondent's recidivism risk that is based on respondent's lack of empathy or remorse is not admissible. However, this does not prevent the State from introducing factual evidence of respondent's lack of remorse and empathy for his victims. Although the State may introduce such factual evidence, it may not link these facts to a conclusion that respondent has a higher recidivism risk based on these facts.

The fourth factor the MDT considered in assessing respondent's recidivism risk was the number of victims he has admittedly abused. At hearing, Dr. Phenix and Dr. Campbell both testified that the Static-99 already includes an

assessment based on number of victims. Essentially, an individual is given a score of 0 if he had no prior charges or convictions relating to sexual offenses, a score of 1 if he had between 1-2 prior charges or 1 prior conviction relating to offenses, a score of 2 if he had between 3-5 charges or 2-3 prior convictions relating to sexual offenses, and a score of 3 if he had 6 or more charges or 4 or more prior convictions relating to sexual offenses. Dr. Phenix and Dr. Campbell testified that, once an assessment has been made using the Static-99, it is improper for the assessor to double-count or give further consideration to factors already addressed by the Static-99. Although Dr. Mart admitted that the Static-99 considers the number of victims an offender has abused, he noted that it makes no distinction between an individual with six victims (or charges) and an individual with twenty or fifty victims. Thus, Dr. Mart testified, he felt that the number of victims in this case should have an impact on respondent's recidivism risk beyond that contained in the Static-99.

Based on the foregoing evidence, the Court finds that expert testimony from the MDT that is based on this factor, apart from and in addition to its inclusion in the Static-99, is inadmissible as it would not be the product of reliable principles or methods. See RSA 516:29-a, I(b). Nevertheless, factual evidence of the number of victims respondent has admitted to abusing is admissible. Furthermore, any of the experts, including Dr. Mart, may testify about the limitations on the Static-99's consideration of this factor. Finally, to the extent it is relevant to prove that respondent "has serious difficulty in controlling his or her behavior so as to pose a potentially serious likelihood of danger to others," see

RSA 135-E:2, VI, expert testimony about how the number of respondent's victims bears on that issue may be admissible.

The fifth factor the MDT considered in assessing respondent's recidivism risk was that respondent is a chronic offender of a predatory nature. Dr. Mart testified that although the "predatory" aspect of this factor is considered by the Static-99, the "chronic" aspect (in terms of number of victims) is not fully considered by the Static-99. However, Dr. Mart's concerns regarding the Static-99's ability to consider the number of respondent's victims is addressed by the Court's Order relative to the fourth factor (that respondent has admitted to victimizing between twenty and fifty individuals). Accordingly, discussion of this concept under the guise of an additional "factor" would amount to double-counting. As previously noted, Dr. Phenix and Dr. Campbell both testified that double-counting is unreliable and reduces the predictive accuracy of the risk assessment. See 516:29-a, I(b) (requiring that testimony be "the product of reliable principles and methods). Accordingly, because it would not comply with RSA 516:29-a, I(b), testimony based on the additional factor that respondent is a "chronic offender of a predatory nature" is not admissible. No expert may testify to an opinion that was generated based on consideration of this factor apart from or in addition to its consideration in the Static-99. Furthermore, no expert may testify that the jury should further consider this factor, beyond its inclusion in the Static-99, in determining respondent's recidivism risk.

The sixth and final factor the MDT considered in assessing respondent's recidivism risk was that all of respondent's victims were male. At hearing, Dr.

Mart acknowledged that the Static-99 addresses the issue of victim gender. However, he noted that the Static-99 only considers whether an individual had any male victims, and does not consider whether an individual had only male victims. Dr. Mart later admitted that there is no empirical evidence that an offender's recidivism risk is increased if his victims were all male. Based on the evidence presented, the Court finds that testimony regarding this factor would be both unreliable, see RSA 516:29-a, I(b), and not based on sufficient facts or data. See RSA 516:29-a, I(a). Accordingly, no expert may testify to an opinion that was generated based on consideration of this factor apart from or in addition to its consideration in the Static-99. Furthermore, no expert may testify that the jury should further consider this factor, beyond its inclusion in the Static-99, in determining respondent's recidivism risk. However, the State may introduce factual evidence that respondent's victims were male.

The Court will now address the admissibility of the MDT's final conclusions regarding respondent's recidivism risk. The MDT reached its final conclusions about respondent's recidivism risk by assessing respondent using the Static-99 original norms, and then considering how the aforementioned additional six factors impact respondent's recidivism risk. Because the Court has previously determined that the original norms do not comply with RSA 516:29-a, members of the MDT may not testify about those norms, nor may they provide testimony that is based on those norms. In addition, the Court has determined that respondent's failure to complete treatment was the only additional factor properly considered by the MDT.

Furthermore, testimony by the MDT regarding how respondent's individual recidivism risk is impacted by his failure to complete treatment would necessarily be based on clinical judgment. Because the experts have testified that clinical judgment is unreliable, the members of the MDT may not testify that respondent's individual recidivism risk is impacted by his failure to complete treatment, and may only testify generally that failure to complete treatment has been linked to increased recidivism rates. See RSA 516:29-a, I(b) (requiring that testimony be the product of reliable principles and methods). In addition, Dr. Campbell and, to a lesser extent, Dr. Phenix, testified that the Static-99 is unreliable when used as a stand-alone instrument to predict a particular offender's recidivism risk, and is therefore not recommended for such use. In other words, according to the experts, an evaluator should not base his or her conclusions about an individual's recidivism risk solely on the Static-99. Accordingly, the members of the MDT may not testify about respondent's recidivism risk based solely on his Static-99 score. See RSA 516:29-a, I(b) (requiring that testimony be the product of reliable principles and methods). The State may introduce evidence of respondent's score on the Static-99, and the range of recidivism risk associated with that score, but not testimony about respondent's specific placement within that range.

In considering the evidence presented at hearing and arriving at the above conclusions, the Court credits the testimony of Dr. Phenix more heavily than that of Dr. Campbell. Dr. Phenix, the State's expert, wrote the protocol for evaluating sexual offenders in the State of California. She has drafted a proposed methodology for

applying the Static-99 with the "new norms." Significantly, she has testified for both the State and respondents in sexual offender commitment hearings. Her testimony was candid, direct and credible.[5]

In contrast, Dr. Campbell, the respondent's expert, seemed defensive, inconsistent, and biased. He only testifies for respondents in sexual offender commitment hearings and has never recommended that a respondent be committed. In addition, he knowingly associated, for five years, with the publication "Issues in Child Abuse Accusations" after learning that its publisher, Dr. Ralph Underwager, had given an alarmingly pro-pedophilia statement to a European journal. Dr. Campbell's book, entitled Smoke and Mirrors: The Devastating Effect of False Sexual Abuse Claims, further suggests his slant, as does his extensive reliance during his testimony on another of his own books, which he conceded on cross-examination contained several significant errors. On balance, the Court found the testimony of Dr. Phenix less biased and gave it more weight.

At hearing, the State indicated that the MDT may need to re-evaluate respondent so that it can apply the new norms. Respondent's counsel challenged the MDT's authority to re-evaluate respondent. The Court need not decide that issue, however, because the development of the new norms does not necessitate re-evaluation in this case. Because the test component of the Static-

[5] The Court notes that, by pleading filed March 13, 2009, respondent has attempted improperly to impeach Dr. Phenix's credibility by discussing her testimony in an unrelated Missouri case. As the State correctly argues, respondent's window of opportunity for such impeachment closed at the end of the Daubert hearing, and it is unfair for respondent to attack Dr. Phenix by way of a pleading to which she is unable to respond. Without such a response by Dr. Phenix, the Court has no way to independently assess these allegations. Accordingly, the Court will not consider this evidence in assessing Dr. Phenix's credibility.

99 has not changed since the MDT evaluated respondent, members of the MDT may testify about respondent's score on the Static-99. They may then testify about the corresponding ranges of recidivism risk set forth in the new norms. Because the Court will not permit testimony about where respondent would fall within that range, the MDT does not need to re-evaluate respondent in order to testify about all aspects of the Static-99 the Court has deemed admissible.

Respondent argues in his pleadings that the "more searching inquiry" involved in State v. Cressey applies to the evidence at issue in this case. See 137 N.H. 402, 408-10 (1993); Baker Valley, 148 N.H. at 615. In assessing the reliability of "controversial child psychology and repressed memory expert testimony," the Court in State v. Cressey, 137 N.H. 402, 408-10 (1993), conducted a more searching inquiry than that conducted under Daubert. Baker Valley, 148 N.H. at 615. The Cressey test

> considers: (1) the presence of objective, quantifiable evaluation results; (2) the existence of a logical nexus between the expert's observations and conclusions; (3) the verifiability of any interpretive steps; and (4) the likely difficulty of effective cross-examination of the expert. The Cressey test differs from the Daubert standard in that it focuses not only upon the reliability of an expert's methodology, but also the reliability of the conclusions and results of that methodology. Under Daubert, the focus of the reliability analysis must be solely on principles and methodology, not on the conclusions they generate.

Baker Valley, 148 N.H. at 615 (citations, quotations, brackets, and ellipses omitted). Although the Cressey Court found such an analysis appropriate in that case, the New Hampshire Supreme Court has "never used this test in other contexts." Baker Valley, 148 N.H. at 615. In explaining its reasons for conducting this more searching inquiry, the Cressey Court noted several

concerns specific to child psychology and repressed memory expert testimony, including that: (1) because "the evaluations of the children deal[t] almost exclusively in vague psychological profiles and symptoms, and unquantifiable evaluation results," and "standardized tests and readily quantifiable results" were not "part of the psychological evaluations," the testimony was inherently uncertain; (2) the expert's conclusions could not be independently verified because there was no "recognizable, logical nexus between many of the identified symptoms," such as that the children drew males when asked to draw a person, or that the children's drawings of people were missing hands or feet, "and the conclusion that the children have been sexually abused"; (3) much of the information relied on by the expert only "beg[an] to take on some sort of meaning ... when ... interpreted and evaluated by someone in her field of expertise," and the necessity of this interpretive step added "yet another variable to the whole evaluation process"; and (4) a thorough cross-examination might not "effectively expose any unreliable elements or assumptions" in the expert's testimony. Cressey, 137 N.H. at 408-10 (brackets added).

As set forth below, the Cressey standard would not impact the admissibility of any evidence the Court has previously determined to be admissible in this case. Furthermore, although the Cressey inquiry may render the MDT's ultimate determination about respondent's particular recidivism risk inadmissible, the Court has already excluded such testimony on other grounds. Accordingly, the Court need not determine whether Cressey actually applies to such evidence. See State v. Forbes, 157 N.H. 570, 576-77 (2008).

The Court will first address, briefly, why the more searching inquiry conducted under Cressey would not prevent admission of testimony relating to the test component of the Static-99. First, testimony relating to the test component of the Static-99 meets the requirement of "objective, quantifiable evaluation results," see Baker Valley, 148 N.H. at 615, because respondent's score on the Static-99 was calculated based on the presence or absence of objective, static factors. In addition, there is a "logical nexus" between the observations used to score respondent on the test and a finding of increased recidivism, see id., because the test itself is based on empirically-validated factors. Furthermore, to the extent there are any interpretive steps involved in the test component of the Static-99, such interpretation relates only to the determination of whether the objective, static factors set forth in the Static-99 are present in a particular case. Thus, these interpretive steps are verifiable. See id. Finally, it would not be difficult to cross-examine an expert on his application of the test component of the Static-99: if the expert improperly "found" the presence of a particular factor, cross-examination would effectively make that clear to the jury. See id. Thus, even if Cressey applied to testimony about the test component of the Static-99, it would not make such testimony inadmissible.

In addition, the Cressey inquiry would not prevent admission of the Static-99 new norms. Because the Court has determined that an expert may only testify to the new norm recidivism range that corresponds to respondent's score on the Static-99, such testimony meets the requirement of having objective, quantifiable evaluation results. See id. Indeed, the recidivism range is based on

the recidivism of past sexual offenders with that particular Static-99 score. Furthermore, there is a logical nexus between a higher score on the Static-99 and a higher recidivism range, see id., because each factor set forth in the test component of the Static-99 is empirically validated. In addition, the only interpretation required to report the recidivism range involves looking at a table to see what range corresponds to a particular Static-99 score, and this interpretive step is completely verifiable. See id. Finally, if an expert improperly translated respondent's Static-99 score into a recidivism range, such inaccuracy would be readily exposed on cross-examination. See id. Thus, even if Cressey applied to testimony about the Static-99 new norm recidivism range that corresponds to respondent's score on that test, it would not make such testimony inadmissible.

Furthermore, the Cressey inquiry would not prevent admission of testimony about respondent's failure to complete treatment and a link between failure to complete treatment and increased sexual offender recidivism. First, such testimony would be based on objective evaluation results, see id., because respondent either has or has not completed treatment. Furthermore, there is a logical nexus between failure to complete treatment and a conclusion of increased recidivism risk, see id., because this factor has been empirically validated. In addition, the interpretive step involved in this process is verifiable, see id., because the empirical data provides an indication of the link between treatment failure and recidivism, and whether it is an inverse or direct relation. In other words, testimony about whether failure to complete treatment increases or decreases recidivism risk, and how strong that impact is, can be verified by the

empirical data supporting the factor. Finally, if an expert mischaracterizes the link between treatment failure and sexual offender recidivism, that mischaracterization can be effectively addressed on cross-examination. See id. Thus, even if Cressey applied to testimony about the empirically-validated link between failure to complete treatment and increased risk of sexual offender recidivism, it would not make such testimony inadmissible.

Based on the foregoing, application of the Cressey standard would not bar any evidence the Court has deemed admissible in this case, and the Court therefore need not determine whether the Cressey standard actually applies to such evidence. See Forbes, 157 N.H. at 576-77.

Summary of Permissible Testimony Under RSA 516:29-a.

In sum, the Court finds that the following **expert** testimony is admissible at the trial in this matter: testimony regarding the MDT's use of the Static-99 to assess respondent's recidivism risk, and his score (on a 0 to 12 scale) on that test; general testimony regarding the range of recidivism risks that corresponds to respondent's score on the Static-99, as set forth in the new norms; testimony about the Static-99's limited ability to consider the number of respondent's victims; and testimony about the empirically-validated link between failure to complete treatment and increased recidivism rates in sexual offenders.

The following **factual** testimony is also admissible at the trial in this matter: testimony that respondent has never successfully completed treatment for his sexual behavior; testimony that respondent has not shown remorse for his actions; testimony that respondent has not shown empathy for his victims;

testimony regarding the number of victims respondent has admittedly abused; and testimony that respondent's victims have been male. This is not an exhaustive list of permissible factual testimony.

Pursuant to RSA 516:29-a, the Court will not permit the following forms of **expert** testimony: testimony that the jury should consider as additional "factors" that respondent has not shown remorse for his actions, has not shown empathy for his victims, that respondent is a "chronic offender of a predatory nature," or that respondent's victims were entirely male; testimony that the jury should give a particular amount of additional weight to the number of respondent's victims due to the Static-99's limited consideration of that factor; testimony attempting to quantify the impact of respondent's failure to complete treatment on his recidivism risk; testimony attempting to place respondent at a particular point within the ranges of recidivism risk outlined by the new norms; testimony that respondent's recidivism risk is necessarily within the ranges outlined by the new norms; and testimony attempting to assign respondent either a quantified (for example, percentage) or a qualified (for example, words like high, medium or low) recidivism risk.

SO ORDERED.

4/28/09
Date

Gillian L. Abramson
Presiding Justice